DSTO

Technical and Operational Tender Evaluations for Complex Military Systems

Andrew P. Gabb and
Derek E. Henderson

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Technical and Operational Tender Evaluations for Complex Military Systems

Andrew P. Gabb and Derek E. Henderson

Information Technology Division
Electronics and Surveillance Research Laboratory

DSTO-TR-0303

ABSTRACT

Technical Report

This report examines technical and operational tender evaluations for complex computer based military systems. Detailed recommendations are made with respect to getting the right technical proposals, developing an evaluation model and controlling the evaluation process.

APPROVED FOR PUBLIC RELEASE

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Published by

DSTO Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury South Australia 5108

Telephone: (08) 259 5296
Fax: (08) 259 5980

AR-009-625
April 1996

APPROVED FOR PUBLIC RELEASE

Executive summary

Introduction

This report examines technical and operational tender evaluations for complex computer based military systems. Detailed recommendations are made with respect to getting the right technical proposals, developing an evaluation model and controlling the evaluation process. This report is completely derived from *A Review of Navy's Technical and Operational Evaluation Practices* (DSTO-TR-0193) by the same authors, with modifications to allow public release.

The evaluations under review are evaluations of responses to a Request for Proposal (RFP) or Request for Tender (RFT) for complex computer based systems, such as combat systems, information systems and ship control and management systems. The main objective of the evaluation as a whole is selection of the supplier of the required systems. The objective of the technical and operational evaluation is the assessment of the suitability of the technical proposals against the requirements defined in the functional and performance specifications in the RFP/RFT, and the determination of the relative ranking of those offers on technical and operational grounds.

It should also be noted that our review concentrated mainly on the technical evaluation of computer based systems, particularly combat systems. Despite this, we believe that many of the findings and recommendations of this review will be useful in improving evaluations in other technical and assessment areas.

Improving evaluations

The quality of an evaluation will be directly affected by the quality of the specifications, the quality of the tendered information and the quality of the evaluation model and process.

Although it is not necessary for the evaluation model to exactly follow the structure of the specification, a poor specification will make the development of the evaluation model very difficult. The development of specifications has also been investigated by the authors, and specifications following these guidelines (DSTO-TR-0192 *Navy Specification Study - Report 3 - Requirements and Specifications*) will generally lead to an acceptable evaluation tree.

The tendered technical proposal is one of the main inputs for the evaluation process, and the structure and content of these proposals are important factors in the effectiveness and efficiency of evaluations. Acquirers have limited control over the quality of the tenders and the technical proposals therein, but can control this to some extent by appropriate directions to tenderers in their preparation of tenders. Acquirers can influence the quality of tenders by defining what the tender should address, how the tenderer should address it and by exposing the tenderers to the evaluation process, showing why the information is needed, and identifying the risks associated with certain

practices. Detailed suggestions are made on how this might be achieved, and how the amount of tendered information might be reduced.

With regard to the evaluation model and process, the following suggestions are made:

- Developing an evaluation model including accommodation for weights and ratings, variants and options, and non-mandatory requirements.
- Preparing for the evaluation including tool selection, and defining the roles and specific guidance for assessors.
- Preliminary review and screening.
- Addressing risk.
- Moderation and collation of assessment results.
- Preparing the evaluation report.
- Ensuring adequate resources are available, particularly for tool support.
- Providing information on the evaluation process to tenderers.
- Recording and using lessons learnt.

Numerical methods

We have found that numerical methods have inherent, unavoidable weaknesses when used for the evaluation of complex systems. They do however have merit when used to provide a separate and secondary "automatic" viewpoint of the evaluation. Advice is provided on the applicability and use of numerical methods.

Evaluation tools

Custom developed database tools have been used successfully in some evaluations, but there have also been serious criticisms about limitations in these tools, including their usability. We reviewed several commercially available tools and found that they fell far short of the custom developed tools. This report recommends that a single database tool be used for all evaluations within an organisation, and outlines requirements for such a tool.

Conclusions and recommendations

The effectiveness and efficiency of technical evaluations can be significantly improved by the application of appropriate processes and tools, and by improvement of these processes based on the lessons learnt from evaluation activities. In particular, guidance needs to be provided on how to conduct an evaluation, and all participants need detailed guidance on their roles in the evaluation. This report provides suggestions regarding what an evaluation process should include, and provides recommendations with regard to the use of tools in evaluations.

Specific recommendations are as follows:

- Establish a defined and monitored evaluation process based on the findings and recommendations of this study.
- Consider the development of a general purpose evaluation database tool to be used in technical and other evaluation areas.
- Dedicated resources be planned and provided for the management of the evaluation tools during evaluations.

- Emphasis be placed on providing guidance to assessors on how to assess tenders against evaluation criteria.
- Guidelines be established with regard to advice to tenderers on the content and format of their technical proposals.
- Use numerically based evaluation methods only as a check against qualitative methods.

Authors

Andrew P. Gabb
Information Technology Division

Andrew Gabb *has been with the Australian Defence Science and Technology Organisation since graduating in Electrical Engineering and Computing Science at the University of Adelaide in 1971. In that time he has acted as a consultant to numerous major projects in the Australian Department of Defence, in the areas of systems and software engineering primarily for combat systems. He is currently a Principal Research Scientist conducting research into C3I front end systems engineering and procurement methods.*

Derek E. Henderson
Information Technology Division

Derek E. Henderson *has been with the Australian Defence Science and Technology Organisation (DSTO) since 1991. Prior to that he was with the Ministry of Defence (MOD) in the UK after graduating in Electronic Systems Engineering at the Royal Military College of Science, Shrivenham, UK. During his time with the MOD he was a member of a project team managing the procurement of software for a submarine command system. Since joining DSTO he has undertaken studies for the Australian Department of Defence in the areas of specification, interoperability, and acquisition of information systems. His current research interests include C3I front end systems engineering, evolutionary acquisition and C3I database systems.*

Contents

Abbreviations

ADO	Australian Defence Organisation
AHP	Analytical Hierarchical Process
AII	Australian Industry Involvement
APET	All Purpose Evaluation Tool
CCTA	Government Centre for Information Systems (UK)
CEPMAN	The Capital Equipment Procurement Manual
CPG	Commonwealth Procurement Guideline
CSU	Consultancy Services Unit
CSV	Comma Separated Variable (format)
DSTO	Defence Science and Technology Organisation
ILS	Integrated Logistics Support
MAUT	Multiattribute Utility Theory
RFP	Request for Proposal
RFT	Request for Tender
SER	Source Evaluation Report
TAAATS	The Australian Advanced Air Traffic System
TDS	Tactical Data Subsystem
TEB	Tender Evaluation Board
TEP	Tender Evaluation Plan
TEWG	Tender Evaluation Working Group
TLS	Top Level Specification

1. Introduction

1.1 General

This report examines technical and operational tender evaluations for complex computer based military systems. Detailed recommendations are made with respect to getting the right technical proposals, developing an evaluation model and controlling the evaluation process. The study was undertaken for the Royal Australian Navy and carried out under DSTO Task NAV 93/067 [Gabb and Henderson 1995e]. This report is completely derived from Technical Report DSTO-TR-0193, *A Review of Navy's Technical and Operational Evaluation Practices* [Gabb and Henderson 1995d], with modifications to allow Public Release.

Typically, computer based systems are difficult to evaluate because of the large numbers of functions which are required and the high level of interaction between different functions. Operational computer based systems include combat systems, information systems and ship control and management systems. Many of these systems have critical real-time requirements, and a very large software component (some Navy combat systems include millions of lines of code). These factors increase the difficulty and complexity of the evaluation.

1.2 Scope

The evaluations under review are evaluations of responses to a Request for Proposal (RFP) or Request for Tender (RFT) for complex computer based systems. The main objective of the evaluation as a whole is selection of the supplier of the required systems. The objective of the *technical and operational* evaluation is the assessment of the suitability of the technical proposals against the requirements defined in the functional and performance specifications in the RFP/RFT, and the determination of the relative ranking of those offers on technical and operational grounds. Technical evaluators will also contribute to other aspects such as the risk perceived in each offer, and each tenderer's ability to supply the systems proposed.

This report distinguishes between assessment and evaluation. *Assessment* is used to mean the assessment of different tenders against one or more individual evaluation criteria. *Evaluation* is used to refer to the comparison between different tenders, and the evaluation activity as a whole.

It is recognised that the technical and operational evaluation is only part of a larger evaluation process. Other assessment areas which will normally also be important to an evaluation in the Australian Defence Organisation (ADO) include [CEPMAN 1995]:

- Contractual
- Project management
- Australian Industry Involvement (AII)

- Integrated Logistics Support (ILS)
- Financial aspects.

Although these areas are not addressed in detail in this report, their evaluation would benefit from following the same principles as for evaluating the technical and operational aspects. Such a common evaluation framework would help ensure a coherent and integrated approach to all facets of tender evaluation.

1.3 Background

Technical evaluations are both difficult and time consuming. They require a large number of technical and operational personnel spread over numerous diverse technical areas for an extended period of time, requiring skilled and dedicated coordination. They are typically conducted with very tight time constraints, and the requirements for ensuring fairness and confidentiality further add to the stress of the activity.

Evaluations for large systems (e.g. a ship) will typically involve reviewing hundreds of documents. The greater the number of tenderers competing, the greater the magnitude of the task. Information relating to a particular technical area or subsystem (e.g. the ship's combat system) may be spread throughout a high percentage of the total documentation, all of which will need to be reviewed.

Evaluations are made more difficult when a number of variants and options are proposed. This has the effect of increasing the number of system configurations which have to be evaluated. In addition, if individual tenderers offer options for several subsystems then these also have to be evaluated, not only as stand alone subsystems to establish their inherent performance capability, but also as part of the overall system to establish their level of integration.

1.4 Policy Context

This paper has been written within the context of the policy defined for the Australian Defence Organisation (ADO). While an understanding of this policy is not critical to an understanding of this paper, the policy lays down the framework within which our study was based, and establishes the terminology which is used throughout the paper.

1.4.1 CEPMAN

The definitive policy regarding evaluations in the ADO is contained in CEPMAN, *The Capital Equipment Procurement Manual* [1995]. Relevant sections include:

- Part 4, Chapter 4: Request for Tender and the Tender Evaluation Plan.
- Part 4, Chapter 5: Tender Evaluation and Source Selection.
- Part 3, Chapter 16: Project Risk Management.

With regard to technical evaluations, CEPMAN emphasises the following:

- Tenders must be evaluated against the criteria identified in the RFT, the tenderers' ability to supply and the determination of risk.
- Communication with the tenderers throughout the evaluation process is restricted to clarification of issues relevant to the evaluation process. Requests for clarification are approved by the Tender Evaluation Board (TEB) Chair.

It also explains techniques for reducing the number of tenders evaluated in detail by:

- Screening - eliminating tenders on the basis of incompleteness.
- Shortlisting - eliminating tenders which are clearly not competitive.
- Setting aside - setting aside tenders which appear not to be competitive but which may be reconsidered if other tenders reveal serious shortcomings during detailed evaluation.

CEPMAN also addresses many of the higher level evaluation issues including:

- The responsibilities of the Project Manager and the formation of the Tender Evaluation Board (TEB) and Tender Evaluation Working Groups (TEWGs).
- The concept of value for money.
- The need for probity, confidentiality, ethics and fair dealing.
- The development of the Tender Evaluation Plan (TEP).
- The contents, format and processing of the Source Evaluation Report (SER).

While source selection is the main objective of evaluations, the SER is the main output of the evaluation process. CEPMAN provides comprehensive advice on how the SER should be structured, and what it should, and should not, contain. CEPMAN states:

> The SER must:
>
> a. State the level of compliance of all tenders with respect to the RFT requirements.
> b. Cover all options arising out of the evaluation of tenders.
> c. Provide a recommendation regarding discrimination between tenders against the evaluation criteria and provide a basis for the recommendation of a preferred source.

It also provides the critical advice that the Tender Evaluation Plan (TEP) should be based on the SER requirements.

With regard to detailed evaluations, each area of non-compliance should be assessed as to the impact of non-compliance. Similarly, where a tender exceeds the requirements, the benefits need to be assessed.

Risk assessment of each evaluation area is required, assessing the risk to capability, cost and project schedule. CEPMAN provides a good basis for understanding risk issues in Part 3, Chapter 16, *Project Risk Management*. Risk assessment is discussed further in section 6.5.

CEPMAN avoids detailed guidance on scoring methods and evaluation tools on the basis that no single method or tool will be applicable in all circumstances.

It also states that the SER must be based on "a qualitative narrative supporting the tender evaluation findings" and that "project managers must not conduct an evaluation on numeric values alone". It warns that where a numerical method has been used as the basis for the findings, the method should be supported by a sensitivity analysis to show that differences in scores genuinely reflect differences in capability. Numerical methods are discussed in section 5.4.

1.4.2 Commonwealth Procurement Guidelines (CPGs)

The Commonwealth Procurement Guidelines (CPGs) are issued by the Department of Administrative Services, and are a series of papers on procurement policy and professional practice. CPGs 1 and 8 are referred to by CEPMAN which is based on the CPGs. While CEPMAN probably has most of the relevant information and guidelines needed by Project and support staff, the Guidelines are readable and informative in a general sense, and we found them useful as a background to CEPMAN.

CPG 1 *Getting Value for Money* "is designed to assist management, procurement staff and end users to buy well without unnecessary costs". It addresses the concept of value for money in relation to the steps necessary in the buying process.

Table 1. *Risks in evaluations*

What can go wrong?	Likely consequences	How to deal with them
Inappropriate supplier selected	Supplier proves unacceptable Supplier unable to fulfil contract Supplier not financially viable	Perform financial and technical check on supplier before awarding the contract Reject offers from unacceptable suppliers Improve evaluation procedures
Inappropriate product selected	Product offered does not meet need	Ensure users are involved with the evaluation Improve technical evaluation procedures
Formal evaluation procedures not observed	Inconsistent evaluation of offers Potential for ethical problems Subjective evaluation outcome	Perform regular audits of procedures Ensure staff are suitably trained and experienced
Breaches of security	Claims of unethical or unfair practices Political intervention Damage to national security	Maintain formal security procedures Perform regular security audits and reviews

CPG 8 *Managing Risks in Procurement* "outlines the philosophy, principles and practices for managing procurement risk". Unlike CEPMAN it not only addresses the risks in proceeding into a contract with a tenderer, but also addresses the risks in the procurement process itself. In particular it identifies the evaluation and source selection activities as being the highest risk in the entire process, with risks of various aspects of these activities ranging from high

to extreme. Table 1, reproduced from this Guideline, shows the risks associated with evaluating offers.

1.5 Costs of tendering industry survey

Recently a survey was commissioned to review Defence procurement practices, particularly with regards to costs of industry tendering for Defence projects. The results of the survey, which encompassed 80 firms, were published in *Costs of Tendering Industry Survey* [1994].

Respondents were critical both of the detail required in tenders, and what they perceived as long evaluation and decision making periods. Reasons given for lack of success included:

- Criteria given were neither clear nor prioritised.
- Lack of expertise in the evaluation team.
- Inability to assess tenders correctly.
- Poor understanding of value for money.

These criticisms might be dismissed as typical of unsuccessful tenderers. They do, however, show that there is a need to make the evaluation activity as efficient as possible, and to have a clearly defined process which can assure both tenderers and Defence decision makers of the effectiveness and objectivity of evaluations.

2. The objectives of evaluations

This section examines the requirements driving the evaluation activity: the objectives of the evaluations, the information on which the evaluation is based (the inputs) and the results of the evaluation (the outputs).

2.1 Objectives

The main objective of the evaluation activity as a whole is the identification of the preferred tenderer. Source selection should be based on the offer which provides the "best value" [CEPMAN 1995; CPG 1 1989; Millett 1994].

For technical and operational evaluations, the main objective is to provide an assessment of the following:

- The level of compliance and performance of each offer.
- The risk perceived in each offer.
- A ranking of offers on the basis of the above.
- Qualitative evidence supporting the above.

A secondary objective of the evaluation is to identify areas in each tender which will require further action during contract negotiation, and to comment on what action might be required. While this may be seen to be an activity which might be carried out more efficiently after source selection (on the basis that only one proposal needs to be considered) it is to some extent a natural byproduct of the

evaluation process. Each proposal is considered in great detail during the evaluation, and non-compliances need to be assessed as to the possibility and impact of their rectification during contract negotiation, as an evaluation issue. Recording this information will generally require very little additional effort.

It should be noted that the availability of this information does not remove the need to review offers prior to contract negotiation, but it does increase the efficiency and quality of that review.

While our study showed that few personnel enjoy the evaluation activities, which tend to be viewed as a trial of judgment, maturity and endurance, there are also several benefits in staff development arising from being part of the evaluation team:

- They gain a good exposure to the capability of industry.
- They see different approaches to the same problem, leading to a broader experience base.
- They often gain an understanding of the value of good requirements and the danger of poor requirements.

2.2 The effect of performance based specifications

The emphasis placed on the use of performance based specifications for Defence projects will have some effect on the evaluation process. The use of performance based specifications can increase risk in the acquisition of complex systems [Gabb and Henderson 1995], particularly when the system supplier is inexperienced or eager to exploit the lack of detail in requirements. Under these circumstances, the contractor's ability and past performance becomes a key part of source selection [Millett 1994].

Such specifications can also make the assessment of "best value" easier [Millett 1994], because they focus on the capabilities to be provided rather than the solution, i.e. the benefits are easier to compare against the costs in a cost benefit analysis.

However, while the requirements (and hence the evaluation criteria) will be performance based, the technical proposals will show how the tenderer intends to provide the required functions and capabilities. Under these circumstances, the actual assessment of the tenders against the criteria will be more difficult, and more prone to risk, because of the need to determine if a proposed implementation is likely to provide the required functions and performance [Macphee 1992].

2.3 Inputs

The inputs to the technical evaluation process will normally be:

- The requirements for the system itself (from the Statement of Requirement).

- Other requirements for the supply of the system (from the Statement of Requirement and other areas of the Request for Tender, including the draft conditions of contract).
- The tendered information - most of the information required will be contained in the technical proposal.
- Information from other sources regarding the tendered products.
- Guidance from high level policy and the Technical Evaluation Plan (TEP), Board (TEB) and Working Group (TEWG) leaders.

The quality of the evaluation will be dependent on the quality of the inputs, in particular the specifications, the technical proposals and the guidance to assessors, as well as on the quality of the evaluation process.

The quality of specifications is addressed in our review of specification practices [Gabb and Henderson 1995], and is further discussed here in section 3.

The quality of the guidance is dependent both on the evaluation process itself (see section 6) and the way in which the objectives and procedures of the process are communicated to assessors (see section 6.4).

Further discussion of improving the quality of the technical proposals is provided in section 4.

2.4 Outputs

The main output of the evaluation process is the Source Evaluation Report (SER) as discussed in section 1.4.1. It is important that the evaluation plans and process, and other interim and lower level outputs focus on providing this output in a form consistent with the required outputs. In particular the SER requires a qualitative narrative showing how the evaluation findings have been derived.

3. The effectiveness of the specifications

The evaluation can be the first serious test for the specifications, as the technical proposals reveal problems in the completeness and consistency of the specification. Planning for the evaluation can provide the first indication that the specification is incomplete, incorrect and/or inconsistent.

Although it is not necessary for the evaluation model to exactly follow the structure of the specification, a poor specification will make the development of the evaluation tree quite difficult (see section 5.3). Gabb and Henderson [1995c] provide advice on writing good specifications. Generally specifications following these guidelines will lead to an acceptable evaluation tree. Specification practices that are likely to assist in the evaluation activity include [Rushforth et al. 1990]:

- Using a logical and tested structure.
- Specifying only one requirement per clause.

- Collecting requirements addressing the same subject in the same part of the specification.
- Separating and clearly indicating mandatory and non-mandatory requirements. (Separation in this case does not mean that the different types of requirements should be in different parts of the specification, but might involve specifying mandatory requirements before non-mandatory requirements in each subject area, for example.)
- Minimising duplication.
- Avoiding the specification of a requirement partly by performance and partly by solution. Otherwise the tenderers may have difficulty in proposing any reasonable solution which will meet all the relevant requirements.

4. Getting the right technical proposals

As one of the main inputs for the evaluation process, the structure and content of tenders are important factors in the effectiveness and efficiency of evaluations. Acquirers generally have limited control over the quality of the tenders and the technical proposals therein, but can control this to some extent by appropriate directions to tenderers in their preparation of tenders. Projects can influence the quality of tenders in three ways:

- By stating what the tender should address.
- By stating how the tenderer should address it.
- By exposing the tenderers to the evaluation process, showing why the information is needed, and identifying the risks associated with certain practices.

While formally outside the scope of this study, we believe it is useful to address these issues.

4.1 Problems in tendered documentation

Typical problems with the tendered documentation are as follows:

- The tenders are very variable both in quality and in the honesty of the responses.
- Information relevant to specific criteria is often difficult to find, because of poor cross references and the fragmentation of information through the technical proposal.
- The structure of the technical proposal is often counter-intuitive, due to the collection of information in different formats from different sources.
- The precedence of the different components of the proposal is not stated, leading to difficulties in resolving problems of inconsistency.
- There are often inconsistencies and contradictions in the tendered material. In some cases large amounts of material are duplicated in a different format or with slight variations.
- Much of the information provided is not needed for the evaluation (even though it may have been requested) further cluttering the proposal.

- It is not clear in many cases what material is formally provided as part of the offer (and will be used as the basis of the technical Statement of Work in the contract) and what material is provided as background material.
- Whether options are formally offered is often not clearly identified, with words such as "it would also be possible to provide feature X" leaving the assessor unsure as to the status of the offer.
- Documents referenced in the tender are not provided, and there are numerous missing or unreadable pages.
- The information is provided in multiple levels of specification (e.g. A, B, C type specifications) which spreads the technical information pertaining to each feature.

4.2 Requirements for tender technical content

It may appear simplistic to state that the minimum amount of information which needs to be provided in a tender is (a) sufficient information to enable effective evaluation, and (b) sufficient information to enable the contract to be drafted.

In our opinion the amount of information tendered can be reduced significantly by the following measures:

- Coordinating the requests for information provided by different technical and operational areas, and from other evaluation areas, such as Integrated Logistics Support (ILS), preventing requests for related information in different formats.
- Reviewing each request for information against the real needs of the evaluators and assessors. (It should also be noted that inexperienced assessors may not know what they need, and will require guidance.)
- Ensuring that requests for information clearly show what level of detail will be acceptable, so that tenderers are not misled into providing excess information.
- Reducing the amount of information requested for which the tenderer is not likely to be contractually liable, i.e. which will not be included in the Statement of Work in the contract. Typical information in this category includes detailed design or manufacture data, and promotional material.

It is also important, however, that tenderers provide sufficient information for evaluation purposes. One common dilemma for assessors is the situation where tenderers claim compliance but provide no supporting information (see also section 6.4). CCTA recommends the following clause in requests for proposals [CCTA 1993]:

> A simple statement that the requirement will be met is not sufficient. For each requirement providers are requested to:
> - Explain how the requirement will be met.
> - Explain what options are available and state the comparative costs involved.
> - Identify clearly any aspect of meeting the requirement that is not included in the proposed costs.

It is likely that assessors will need more supporting information for assessing some criteria than in others (e.g. navigation performance). These should be

identified when developing the evaluation model and specifically addressed in the RFT.

Advice on the treatment of options in technical proposals could also reduce the amount of documentation and at the same time improve the lot of the assessors.

- Where an option includes many changes to the overall configuration or functions of an offered system, it is preferable that this be shown as changes ("deltas") to a baseline offer, rather than providing an integral description of the variant system or subsystem. This will avoid the necessity of comparing many (in some cases hundreds of) pages to detect the differences.
- Any option offered must be integral to the offer, i.e. how the optional features or components interact with other system functions/components must be shown, or it will be impossible to adequately evaluate the option. In some cases it will also be necessary to show the relationships between options, where there is likely to be an interaction between them.

4.3 Other tender deliverables

Many of the problems in evaluations stem from the difficulty in finding information in the tenders, and the uncertain status of some of the information provided. We recommend that the following be required in all tenders:

a. Volume identification. Each volume should be uniquely identified by a short prominently displayed volume number.

b. List of contents. A comprehensive list of contents should be provided showing which documents are included in which volume (recognising the fact that some volumes will contain several documents and a single document might span several volumes).

c. Statement of technical document precedence and status. In cases of possible conflicts it is important that the precedence and status of documents constituting the formal offer is known and summarised. In addition to the precedence, this information should clearly show why the document is included in the tender and its status, e.g. whether it (or selected parts thereof) is intended to be included in the Statement of Work in the contract, or supporting technical information, option, example or background information. Tenderers should be advised that it will usually not be necessary for all of their technical proposal to be included in the contract, but they should formally state whether what is provided in the tender is formally offered in response to the Statement of Requirement.

d. Configuration information. An overview of the system components, listing identification, supplier (if applicable) and purpose. All optional components should also be addressed.

e. Statement of system component status. A list of each of the significant system components (typically subsystems, equipment and software) showing the development status or maturity of each component at the

time of tendering. It should be shown whether the component is to be newly developed or in-service, or whether the component is to be derived from an in-service component and in this case the extent of modification. Where the component is in-service, these applications should be identified and the number of installations stated.

f. Statement of Compliance. A Statement of Compliance should be provided, including as a minimum:

- The relevant clause numbers of each requirement (and their source, where there is more than one specification).
- The level of compliance.
- Detailed references to the technical proposal, showing precisely where clear evidence of compliance is shown.
- Comments (e.g. addressing why compliance is incomplete and the perceived impact).

The following comments apply specifically to the contents of the Statement of Compliance:

- A template for the Statement of Compliance should be provided as part of the RFT, both in a preferred database format and in a generic database format (Comma Separated Variable (CSV) is one of the most common database interchange formats used and is recommended). Tenderers should be required to submit their Statement of Compliance in either format.
- For the claimed compliance, we recommend that tenderers be provided with more flexibility than merely stating whether their offer is compliant or not. This will allow them to be more accurate in their self assessment. Other categories might include "near compliance", "partial compliance" and "not applicable" (where in the tenderer's opinion an alternative solution removes the need to satisfy a particular requirement).
- References should be explicit, identifying the area or areas of compliance to the nearest page, and accompanied by clarifying comments if necessary. Numerous inadequate references should be sufficient cause to reject a tender.
- All options must be addressed in the Statement of Compliance.

4.4 Format of technical information

In our opinion, technical information should normally be provided in the format considered appropriate by the tenderer. There are several reasons for this:

- Complex systems are difficult to describe in any format. It should be assumed that the tenderer can best explain the system, and best choose the format in which to present it. Forcing a different format may reduce the understandability of the information.
- If the tenderer is obliged to reformat information, this will incur additional tendering costs, with limited benefit to the acquirers.
- Reformatting may result in information being omitted, because it does not fit in the specified format.

In general, if the tenderer provides the information requested, and provides detailed references in the Statement of Compliance or where the information might otherwise be expected, the format of the material should not be a concern.

5. The evaluation model

5.1 Overview

Evaluation of a complex system involves the assessment of tenders against hundreds or even thousands of criteria based on the Statement of Requirement. This daunting task would be impossible without arranging the criteria in some way so that assessments and comparisons can be made on relatively small numbers of criteria at a time. The arrangement of the criteria in a way which shows the relationships between them and their relative importance, is often referred to as the "evaluation model", which provides the main structural framework for the evaluation process.

5.2 The evaluation tree

The most common form of evaluation model takes the form of an *evaluation hierarchy* or *evaluation tree* where the top level of the hierarchy reflects the major capabilities (or subsystems and major component capabilities) required of the system. Each of these capabilities is then decomposed into lower level capabilities (this process is often referred to as functional decomposition). The decomposition is repeated for each level until the criteria at the lowest levels reflect the requirements for the system. An example of an evaluation tree is shown in Figures 1 and 2.

Each criterion in the tree will then be assigned a *weighting factor* which is an indication of the importance of that criterion compared with other criteria in the same branch at the same level. The weighting factors are usually chosen so that those contributing to a particular branch/level add up to 1 or 100, which can make the assignment of factors and the calculation of the scoring easier. Weighting factors should be assigned even when numerical scoring methods are not proposed, as a means of showing the relative importance of related criteria. In these cases a more qualitative weighting scheme might be used using factors such as "critical", "high", "medium" and "low". Such "factors" are still *ordinal* in that they show priority, but they avoid the mathematical significance implied by numerical factors.

During the assessment of individual criteria, each "leaf" (the lowest level criterion in each branch) will be assigned a *rating* indicating the level of compliance, typically either as a score between 0 and 10, or as a ordinal qualitative measure such as "superior", "compliant" or "non-compliant". Ratings for the higher levels can then be determined on the basis of the ratings for the nodes below them. This process can be continued until a rating for the system as a whole is achieved.

5.3 Developing the evaluation tree

For simplicity of description, the preceding discussion implies that the evaluation tree is derived from the top level requirements using functional decomposition until the criteria conveniently decompose into the requirements (and hence the evaluation criteria) for the system. This will rarely if ever be the case. Instead the Statement of Requirement in the RFT, consisting of the performance and functional specifications (and other requirements in some cases, such as engineering requirements) should form the basis for the evaluation tree, if they have been written correctly [Gabb and Henderson 1995].

Poorly written specifications, with an illogical structure, or which do not follow the broad principles of functional decomposition, will be much more difficult to translate into an evaluation tree. In these cases it will be necessary to derive the tree using the requirements only as a guide, and then to trace the requirements into the tree. One of the difficulties that will often be experienced in this activity is that requirements in the specifications may apply to different parts of the evaluation tree, i.e. a single requirement may translate into two or more evaluation criteria.

Regardless of this, the structure of the evaluation tree will rarely be exactly the same as that of the specifications for the following reasons:

a. It will often be useful to combine several requirements into a single criterion when they are closely related. This will reduce the assessment effort (fewer forms will need to be completed) and in some cases ensure a more accurate rating. An example of this might be requirements indicating that 10 different items of information be displayed to the operator. While these might be technically seen as 10 requirements, an assessment of all 10 in the same criterion would be sensible. In addition, the assessor would be able to assess the impact of the combination of displayed items, possibly reducing (or increasing) the rating penalty that might otherwise be imposed.

b. While there will be little duplication in the requirements tree, the evaluation tree will occasionally contain duplicate or linked criteria, possibly with different weights. Although the concept of criteria contributing more than once to an evaluation might intuitively appear to be poor practice (indeed Roetzheim [1991] states that it should not be done), the effect of multiple contributions is corrected by the weighting factors. It is not only acceptable practice - it is the logical way of reflecting the fact that some functions of an integrated system contribute to more than one of the system's capabilities.

c. In many cases high level requirements in the specification will correspond to high level criteria in the evaluation tree. In other cases the high level requirements will be either implicit, or may not be worded as requirements in the specification. Because the rating of these criteria will be completely dependent on the criteria below them, there are no difficulties in handling this situation. One example of such a criterion occurs when there are specific requirements for compatibility. These will

normally be included in the evaluation tree under a general criterion for compatibility, although there may be no corresponding general requirement.

d. There will be additional requirements which may not be in the specifications, because they do not strictly define technical requirements but which are appropriate for the technical and operational evaluation team to address. Examples are the tenderers' understanding of the requirement, and the tenderers' expertise and relevant experience.

Apart from these exceptions, each requirement in the Statement of Requirement must be traceable to criteria in the evaluation tree, and all criteria must be traceable to requirements in either the RFT package or the Tender Evaluation Plan. When this is assured, the evaluation should be based on the criteria in the evaluation tree rather than the specifications themselves (or some other structure) to ensure the consistency and integrity of the evaluation.

Table 2. *Top Level Specification structure*

1. Introduction
2. Mission
Tasks
Task priorities
Operational environment
Area of operations
Concept of operations
3. General operational requirement
Availability for sea
Operational availability
Survivability
Supportability
Sustainability
Medical
4. Combat system requirement
Command, control, communications, computers and intelligence
Command and control
Combat data system
Tactical data links
Navigation
...
Communications
Administration information systems
Intelligence support
Sensors
Weapons
Aviation
Miscellaneous
5. Platform System Requirements

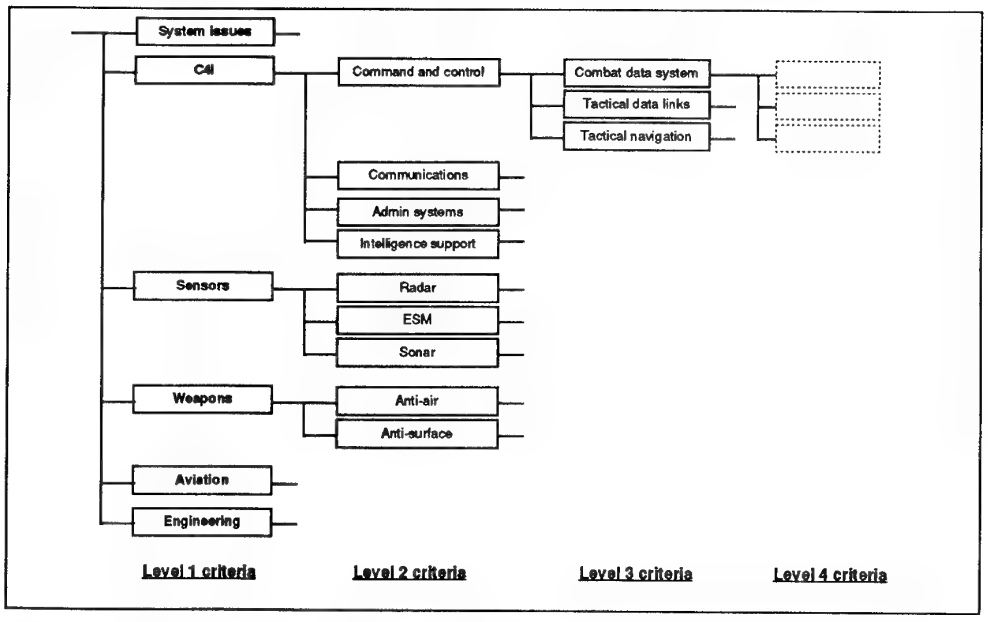
Most combat systems for the Royal Australian Navy, for example, are likely to be defined in a Top Level Specification (TLS) for the ship as a whole. An outline of a TLS is shown in Table 2. This structure is relatively common for complex systems and consists of operational requirements in the earlier sections, and the requirements (generally performance and functional requirements) for major components or subsystems in the later sections. The subsystem requirements should be derived from the operational requirements, reflecting the most

stringent requirements demanded by the different tasks the ship and its combat system are required to carry out.

It should be noted, however, that while such a specification may be based on a functional decomposition of requirements, it is not written in that form. It actually consists of two trees, one based on the operational needs of the ship, the other on the functions and performance of numerous subsystems and major components. Figure 1 shows the top levels of a possible combat system technical evaluation based on the TLS structure. As can be seen, this is based mainly on section 4 of the TLS and addresses the performance of the major subsystems and components. Figure 2 shows a possible branch of the same tree illustrating the lower levels.

After evaluation of the criteria in the tree, the results will be used in a further step, evaluating each tender's ability to meet the operational requirements with regard to the different mission tasks. This step cannot easily be done numerically - each task will draw on different criteria of the tree evaluation, and omit others. Other information may also be drawn from other evaluation trees, such as that for the ship itself. For example, the performance of the Search and Rescue task is unlikely to be dependent on the performance of the ship's weapons, and may require a different radar performance than is required in the ship's primary role. Consequently, this step is likely to be done qualitatively.

Figure 1. *Possible combat system evaluation tree - top levels*

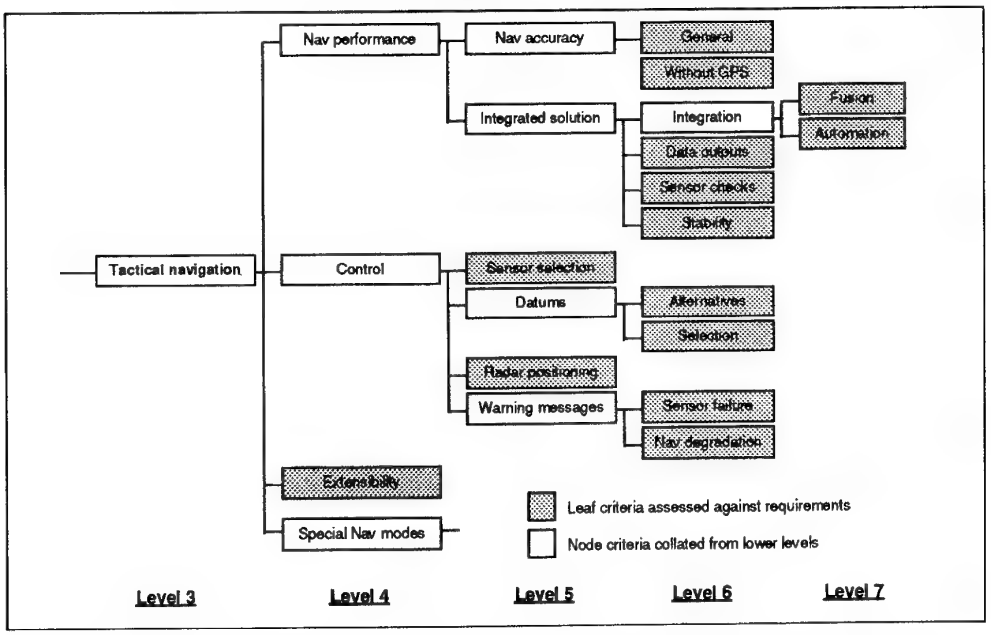


5.4 Numerical evaluation methods

There has been much discussion within the ADO on the value or otherwise of numerical evaluation methods. Many references (including Roetzheim [1991] and CCTA [1990]) propose such methods. Others such as Rushforth et al. [1990] warn of "losing important effects in a sea of numbers", or advise that complex

decisions need properly informed subjective judgment [CPG 1 1989]. CEPMAN [1995] requires a "qualitative narrative" to show the evaluation findings.

***Figure 2.** Tactical navigation branch of evaluation tree*



In our opinion, numerical methods have inherent, unavoidable weaknesses (which are listed at the end of this section) when used for the evaluation of complex systems. They do however have merit when used to provide a separate and secondary "automatic" viewpoint of the evaluation. Used in this way they can provide a check against the primary qualitative evaluation, in some cases revealing flaws either in the weighting or in the qualitative assessments, thereby improving the quality of the technical evaluation. It will not be unusual for the numerical and qualitative evaluations to show different results. It is the analysis of *why* this has occurred which will provide the added value.

For numerical evaluations the assessment of node criteria (as opposed to leaves) is primarily a mathematical calculation, based on the ratings and weighting factors of the criteria contributing to the criterion being assessed.

While there can be a wide variation of how the ratings are assigned, what form the weighting factors take, and hence how the higher levels are assigned, most evaluation systems are similar. Typical examples can be seen in CCTA [1990] and Roetzheim [1991]. In the CCTA model, for example, weighting factors are between 1 and 100 and ratings between 0 and 10. All weighting factors of criteria directly contributing to a higher node add up to 100, so that a factor of 30 (say) implies that the corresponding criterion is considered to be 30% of the importance of the higher criterion, and considered twice as important as a criterion with a factor of 15. To find the rating of a higher criterion, each rating below it is multiplied by its weighting factor, and the resultant scores are added together and divided by 100.

We see the deficiencies in numerical methods as follows:

a. One of the reasons that numerical methods are promoted is that they provide an illusion of objectivity. In fact, numerical methods require subjective judgment in the development of the criteria, the weights and the assessment ratings, and are just as susceptible to bias, accidental or deliberate, as qualitative methods. Because the bias may be hidden in a "sea of numbers", however, it may be more difficult to detect in numerical evaluations.

b. Numerical evaluations also give an illusion of preciseness and correctness, and are less likely to be examined as carefully or challenged for this reason. One of the critical reasons for the source selection for the TAAATS project being questioned was the consideration of issues qualitatively at a high level after they had been included numerically at the lower levels [Macphee 1992]. In effect, a tenderer was penalised twice for the same perceived deficiency.

c. It makes no sense to combine some criteria numerically. For example, consider comparing a requirement for navigation accuracy with a requirement to use the Ada programming language. Trying to provide weights for the corresponding criteria is difficult. Establishing rating values so that, for example, a shortfall of 10% in navigation accuracy is equivalent to only having 80% of the software in Ada may be appropriate, but would be extremely difficult to justify on a general basis.

d. Modern systems provide a high level of integration of functions and capabilities. It is often not possible to write criteria which will reflect this when combined numerically. For example, the requirements for an administrative computing suite may call for both wordprocessor and spreadsheet products. A tender may meet the individual criteria for both products but there are known compatibility problems when the products proposed are used simultaneously. Under a numerical scheme, the higher level ratings will be calculated automatically. Using a qualitative scheme this problem is more likely to be detected when the higher level criterion is assessed.

e. Claims that the suitability of an offer for a complex system can be incorporated in a single number are at best simplistic and at worst misleading.

f. Numerical methods have difficulty in handling other factors which may be included as part of an assessment, e.g. risk, confidence level of the assessment, impact of deficiencies and non-compliances.

5.5 Assessment results

The evaluation model should also identify how the assessment of each criterion will be represented. While this may appear relatively obvious, it is essential, particularly when using computer based tools (which allow less flexibility than a written assessment), that the outputs and guidance cover all eventualities which the assessor may face. The results should encompass:

- Compliance and the impact of non-compliance - whether a tender meets the criterion, and the impact to the project if it does not.
- Performance - how well the proposal meet the criterion. This will normally be merged into a single rating with compliance.
- Risks and impact - whether there are perceived risks in the suggested solution, or the way it is described, and the impact to the project.
- Confidence - in the ratings given, normally based on the amount of evidence available in the proposal or elsewhere.
- Supporting detail - an indication of how well the tenderer has addressed the criterion. This is related to but separate from confidence and can be used later to assess how tenderers responded and improve the evaluation process.
- Justification and references - showing where the evidence was found and the reasons for the assessment ratings.
- Comments on whether there are any negotiation issues arising from the assessment.

How these outputs will be used, and the relationships between them (e.g. that if risk is identified, the impact should also be assessed) should also be determined at an early stage. Because the selection of these outputs could be critical to both the effectiveness and efficiency of the evaluation the outputs should be tested against a set of situations which are likely to be faced by the assessors. Typical exceptional circumstances include:

- The tenderer claims compliance, but evidence contradicts this, or there is no evidence.
- The tenderer assesses the proposal as non-compliant, but it is assessed as compliant in the evaluation.
- The offer appears compliant, but there is solid evidence external to the tender that it will not be, e.g. based on experience with the equipment proposed.
- The criterion is not applicable to this offer, e.g. because the solution provides this functionality in a different way. This is usually an indication that the specification drafter has made incorrect assumptions about potential solutions.
- The offer is assessed as both non-compliant *and* the solution proposed involves risk.
- There is conflicting evidence regarding compliance.
- The proposal is rated as non-compliant, but with low confidence.

5.6 Assigning weights and rating values

There are many different ways in which weighting factors and rating values may be determined. Those discussed here should be viewed as samples. It is important, however, that the method used to assign weights and ratings is carefully defined, documented and consistent across the evaluation model.

5.6.1 Assigning weights

Assigning weights to different criteria will often be the most difficult part of planning the evaluation. Weights reflect the importance of criteria relative to

those in the same branch at the same level. Even if not using a numerical method, it is still necessary to determine the weight which is attached to each criterion, so that assessments of criteria in the tree can be combined in a predictable and repeatable fashion. The weights in such a case might be numerical or use some other ordinal method, as discussed earlier in this section.

Determining weights is not only difficult, particularly when the criteria are not closely related, it is also highly subjective. For this reason, there are likely to be variations between different personnel on what weights should be assigned, and disagreements between personnel with different responsibilities, e.g. between operational and support staff.

As weights are determined, some qualitative description should be provided to show why the weights were chosen [CCTA 1990]. This information is essential if it is necessary to change the weights when, for example, additional criteria are incorporated in the evaluation tree.

One method commonly used to resolve such problems is the Analytical Hierarchical Process (AHP) [Saaty 1980; Tullous and Utecht 1994]. In this method the relative priority of each pair of criteria is assessed by a number of stakeholders and the results combined mathematically to define the weights. The method includes techniques for checking for the consistency of the inputs, so that severe differences of opinion can be identified and addressed.

AHP is used in several proprietary tools and methods including Criterium and APET (see section 7.2).

5.6.2 Assigning rating values

"Rating values" here refers to the meaning given to different ratings, rather than the actual assignment of ratings during the evaluation. Whether the evaluation uses numeric or ordinal qualitative ratings, the rating values need to be carefully defined and communicated to all assessors.

The actual values used, and their meanings, can vary considerably with the evaluation and scoring method used.

One method employed in a recent project used the following ratings for compliance:

- Superior. The tenderer has complied without reservation and offered benefits beyond those sought.
- Compliant. Describes an element of an offer which is assessed as meeting the specified requirements.
- Acceptable. The tenderer has complied without reservation or has offered substantial compliance but with some minor variation which does not substantially alter the requirement.
- Non-compliant. Does not meet the specified requirements.

A possible problem in this approach is that assessors have little scope to show minor differences in compliance or performance between competing offers.

Roetzheim [1991] suggests a numerical rating where a value of 0 is assigned for the minimum acceptable performance and 100 for the best proposed offer. This approach is not always suitable because it assumes that there are no offers with unacceptable features (presumably these have already been discarded), and that the best proposed offer for each criterion is known.

CCTA [1990] defines a numerical rating with a more pragmatic approach as shown in the following table. It should be noted that the CCTA guidance applies to the procurement of information systems.

***Table 3.** CCTA rating values*

Score	Description	Comment
0	No capability	Unacceptable
1	Very limited facility - could be made to meet requirement, but only at a very basic level	May present major difficulties in use
2	Limited facility	
3	Basic facility - not easy to use	
4	Adequate - could be made to meet most of the requirement	Some difficulties in use
5	Fair - not entirely satisfactory	
6	Reasonable - could be made to meet most of the requirement	Difficulties which could be reasonably overcome
7	Fairly good	
8	Good - does not require any manipulation to meet the requirement	Minor problems, easily overcome
9	Very good - for all practical purposes, shortcomings can be ignored	
10	Excellent - cannot be improved upon	

This rating scheme gives the assessor a reasonable flexibility in assigning ratings. While seemingly quite explicit, it is still highly subjective, but this problem can be resolved to some extent by the use of examples.

5.7 Using the evaluation model

The evaluation model will normally not be used until the higher level filtering processes for tenders (involving screening, shortlisting and setting aside - see section 1.4.1) have been completed. All tenders are therefore likely to exhibit a reasonably high level of compliance, at least against the more important criteria.

In the detailed technical and operational evaluation, the assessment activity will establish the compliance rating of each tender against each "leaf" criterion in the evaluation tree. It will normally also involve the assignment of other ratings as shown in section 5.5.

Assessment of the higher level criteria can then be done, collating the results of the criteria below them in the tree.

For numerical evaluations, the compliance rating of higher nodes in the tree can be calculated automatically, but the other ratings will normally be collated by hand. It is common in such cases [e.g. Macphee 1992] not to propagate the

evaluation to the highest level (to produce a single figure of merit) but to address these levels qualitatively, using the numerical figures as a guide.

5.8 Variants, options and common solutions

Alternative solutions in the same tender can either be treated as separate configurations (variants) or as options to a particular proposal. Where there are a large number of options, however, and there is likely to be interaction between them, these approaches will not work easily. If they are treated as variants, there are likely to be a very large number of configurations to be considered. If treated as independent options, it is difficult to assess the performance of combinations of these.

Where two or more tenderers offer identical solutions for a subsystem there is no point in evaluating the common solution more than once. (Evaluators must be certain, of course, that the solutions proposed *are* identical. It is common for different tenderers to offer proposals that are similar, and based on the same components, but not identical. In this case it may be more prudent to consider them as separate solutions.)

It may be useful in some cases to modify the evaluation tree to facilitate the handling of options and/or common solutions. One change may be to try to ensure that all the functionality of each optional/common area is self-contained, in its own branch of the tree. This may need to occur after the evaluation activity has commenced.

Variants, options and common solutions present more of a challenge for the evaluation tools, and are discussed further in 8.3.

5.9 Non-mandatory requirements and excess performance

The foregoing has assumed that all criteria and requirements could be treated identically, with the priority of the requirements reflected by the different weighting factors assigned to the evaluation criteria. Handling non-mandatory requirements, and performance which is beyond that specified, introduces complications.

Non-mandatory requirements are usually indicated by the use of the words "should", "may", or "it is desirable that". They are requirements which are of lower priority than the mandatory ("shall") requirements and are often used to indicate a preference, or a capability of lower benefit than the others. To be compliant, a proposal only needs to satisfy the mandatory criteria. It follows therefore that a numerical assessment of *compliance* should only include mandatory criteria.

Non-mandatory criteria need to be in the evaluation tree, so that they can be assessed at the same time as the related mandatory criteria. They will normally have no weighting, however, with regard to compliance.

An assessment of overall *performance*, however, should include both mandatory and non-mandatory criteria and reflect excess performance. Where a numerical assessment of performance is required, both types of requirements will require weightings.

6. The evaluation process

6.1 Overview

It is now recognised that it is difficult to consistently produce quality products without a defined systematic process. Similarly, improvement of quality is extremely difficult to achieve without regular review of the process, and evaluation of the resultant products, with an aim towards process improvement. This also applies to tender evaluation, where the product of the evaluation is the source selection and its justification.

This section addresses the evaluation process and its products with an aim to the improvement of that process. Rushforth et al. [1990] emphasise the importance of "standardising the approach to evaluation and in ensuring that satisfactory standards of execution are maintained", stating that such a process meets the following needs:

- The need to have an objective, overt and neutral ranking process.
- The need to demonstrate fairness.
- The principle that criteria must be agreed before tenders are examined (as also required by CEPMAN).
- The need for structures and checklists which are understandable and usable by the project teams undertaking evaluations.

It should be noted, moreover, that the recommended management of the risks of inadequate evaluation suggested by CPG 8 [1992] is achieved by procedural means (see section 1.4.2).

This report does not attempt to define the evaluation process in detail. Instead it examines the issues that such a process should address, and suggests some of the activities and guidance that the process should contain.

6.2 Preparation

6.2.1 Planning

Planning for the evaluation as a whole includes developing a Tender Evaluation Plan (TEP) as called for and described in CEPMAN (see section 1.4.1). Planning for the technical and operational evaluation must include the following steps:

- Developing the evaluation model.
- Selecting a tool or tools to support the model, identifying changes that need to be made and making those changes.

- Providing guidance to the evaluation teams on the evaluation model and the use of the tools.
- Determining the Technical Evaluation Working Groups (TEWGs), and the personnel who will be involved.
- Allocating criteria in the evaluation model firstly to TEWGs, and secondly to assessors.
- Planning other resources.

Having some understanding of the types of solutions which are likely to be proposed by tenderers may affect early planning and reduce problems later in the evaluation process. Providing a draft release of the specifications to potential tenderers [Gabb and Henderson 1995] and allowing tenderers to present their likely solutions at an early stage would be very useful in this regard, but only if those responsible for the relevant part of the technical evaluation attend.

The plan should be focused on providing the required inputs for the Source Evaluation Report (SER - see section 1.4.1). This will require that there is general agreement as to what the SER will contain early in the planning process. Outputs from the evaluation process are discussed in section 2.4.

6.2.2 Preparing the evaluation model

The evaluation model should be created in the late stages of the specification development [CCTA 1990]. Although this may seem to be unnecessary, with the evaluation still 6 to 12 months distant, it is important that it be done at this time for two reasons:

- As described in section 3, developing the evaluation model can reveal critical structural and content weaknesses in the specification, which can be rectified at this stage, but rarely later. In addition, considering how requirements will be evaluated can show weaknesses in the expression of those requirements. Hence developing an evaluation model can improve the quality of the specification [Rushforth et al. 1990].
- The RFT will contain an indication of the high level criteria to be used in the evaluation, and possibly some indication of the importance of those criteria. Developing an evaluation model ensures that this high level model is correct, and consistent with the needs of the evaluation [CCTA 1993].

The development of the evaluation model should produce the following [CCTA 1990]:

- The evaluation tree, with associated criteria defined, showing traceability to the requirements.
- Notes showing how the weights (or priorities) were defined.
- Annotations to the criteria to assist in assessments.

6.2.3 Selecting and preparing the tools

The tools must be suitable for the evaluation model and process. If the tools are not matched to the model and process, it is likely that this will cause serious

problems during evaluation, affecting both the efficiency and effectiveness of the evaluation. Tools are further addressed in section 8.

6.2.4 Preparing guidance

Comprehensive and authoritative guidance is needed both to ensure assessors understand their roles, and to ensure consistency across the entire evaluation.

Written guidance should clearly show what is expected of evaluators and assessors, and cover most contingencies which are likely to arise during the evaluation. Numerous examples should be included, and be drawn from the diverse areas being covered by the evaluation. Guidance should include:

- An overview of the evaluation model, showing how different functional areas interact.
- A clear indication of what ratings are required, under what circumstances they are applicable and what each individual rating value means.
- Guidance on how to enter the text fields (such as references, comments, negotiation notes), showing what is required in each field, its scope and applicability and how it will be used in the evaluation. If it is important to use a particular style of wording, this should be clearly stated and examples provided.
- Guidance should be provided on handling exceptional circumstances, such as when and how to apply "benefit of the doubt" (see section 5.5).
- The constraints placed on the evaluation teams should be clearly stated. These should include constraints identified which apply to ethics, fairness and confidentiality. (Those responsible for the evaluation should recognise the fact that few of those in the evaluation teams will have read the higher level policy.)

6.2.5 Selecting the evaluation participants

Prior to the evaluation it is necessary to assign the responsibility for the different criteria to the Technical Evaluation Working Groups (TEWGs) and then, if necessary, to evaluation teams within the TEWGs. For some criteria assessors from more than one team or TEWG will contribute to the assessment, and this also needs to be decided. TEWG and/or team leaders will also need to determine the moderators and individual assessors.

Assessments will normally be most effective if the same assessors are used for the same criteria across all tenders. This will allow them to form a balanced view of all tenders with regard to these aspects. It will also reduce the need for formal instructions for each technical area which might be required to ensure consistency if different assessors are assessing the same area.

To provide further balance, it is advisable that at least two assessors work as a team in these assessments, reducing the likelihood of personal bias or error. While it might be thought that moderation can provide this balance (see section 6.6), the moderator may have limited expertise in some specialised areas.

6.2.6 Selecting the source for assistance

Finally, the plans and the guidance should identify a point of contact which can provide authoritative advice with regard to the evaluation process, to resolve the inevitable uncertainties when they arise.

6.3 Preliminary review and screening

It is preferable that the evaluation team has access to *all* the relevant material at the time of assessment. Assessment of different tenders over an extended period will reduce their ability to compare offers, as will the discovery of additional information at a later time.

For this latter reason, the first step in the technical evaluation needs to be the review and collation of tendered information to identify what has been offered in terms of configurations and options, and where the relevant information can be found. This may take several days and needs to be done by personnel with a system view of the project, who can recognise the significance of different technical areas. This initial review and collation process will also provide some indication of the quality of the offers, and will be the first step in providing advice to the TEB with regard to screening, shortlisting and setting aside of tenders.

The length of the initial review period will depend on several factors including the complexity of the application, the number and quality of the tenders (see section 4), and the number of variants and options offered. In some cases it may require weeks rather than days.

At this stage, before any detailed assessments have been done, it will also be useful to review the evaluation model and the supporting tools. The exposure to the tenders will often indicate that changes at the lower levels may be warranted, usually as a result of innovative solutions or areas where most of the tenderers are having difficulties in meeting the requirements. Examples include the splitting or merging of criteria, their rearrangement to provide more efficient handling of options (see section 5.8), or accommodation of implicit requirements.

6.4 Assessment guidance

Some specific advice which may be considered in the assessment of criteria is as follows:

- Do not take statements in the tender at face value. Unfortunately it is relatively common for tenderers to word technical proposals in ways which encourage misinterpretation.
- Avoid emotive or superlative statements in the justification for the assessment - they may be seen as evidence of bias, or be used out of context.
- Beware of preferring or rejecting well known solutions over those which the assessor is less well acquainted. This is the "devil you know"

syndrome and can apply in either direction. We are aware of an assessment, for example, where a known solution was initially ranked last (out of 4) because the assessors were aware of all of its deficiencies, and little information had been provided on the alternatives. When further specialist advice was sought, the known solution was ranked second.

- If you are not capable of evaluating a requirement to the required level of confidence, because of limited experience or knowledge, say so.
- Stay within your area of expertise and your defined area of responsibility.
- Recognise opposing schools of thought when assessing controversial criteria. One example of this might be in assessing the advantages of using the C++ programming language over Ada. In such areas, there are many who hold strong opinions - the assessment will require additional effort and the justification will require careful wording.
- The information on which a proposal is assessed is not restricted to the tendered information. Other information sources might include, for example, experience with the proposed equipment in Defence, visits to a tenderer's premises, or published assessments of commercial software. It is important to ensure, however, that such information is valid and objective, and applies to the same version of the supplies tendered.
- Ensure that information used to assess a proposal is being formally offered. Material which is an example of what *could* be provided does not provide a firm basis for assessment, for instance.

Two problems often faced by assessors are as follows.

If a tenderer claims compliance, but provides no information to support this claim, the criterion may be difficult to assess. If another tenderer has provided detailed information, and has been assessed as deficient, the problem becomes more serious. There is a risk of rewarding a tenderer who has provided insufficient information, and penalising a tenderer for supplying adequate information. This situation is clearly unfair.

We believe the approach to this situation should be as follows. If the criterion is regarded as important, further information should be sought from the tenderer. If there are numerous similar instances, the tender should be set aside. Otherwise the assessor must estimate the risk in proceeding with limited information, and assign this risk rating to the assessment of that criterion. In the special case that another tenderer has been scored down on the basis of detailed information, we believe that the first tenderer should also be scored down to the same level, and this fact should be noted in the justification for the assessment.

There are some cases, of course, where simply claiming compliance will be acceptable. Requirements that certain explicit data will be recorded, or that a particular standard will be met, for example, may not require clarifying information or additional evidence.

A second case is where a tenderer claims compliance against a requirement but there is clear evidence that the offer is not compliant. This happens too often to seek clarification for each case, and is more prevalent in some tenders than in others. It is important, obviously, for the assessor and moderator to be *sure* that

the assessment is correct, i.e. that the proposal is not compliant in their expert opinion. Special attention needs to be given to the assessment under these circumstances. It may also be useful to try to understand the basis on which the tenderer has claimed compliance - this may have resulted from different assumptions about the environmental conditions under which the performance is required, for example. Assessors should also be made aware of the fact that this situation occurs regularly, and that the self assessments by the tenderer should be treated with caution.

Where this occurs frequently in a particular tender, a note of this should be made in the evaluation report, and consideration should be given to increasing the risk assessment for this tenderer.

Some recent projects have prohibited any form of handwritten notes in tenders. In our opinion, this may detract from the effectiveness and efficiency of the evaluation. Marginal notes can be very useful in drawing the attention of both the marker and other assessors of potential problems and other critical issues in the documents, such as deceptive wording. In this case they act as a warning to the reader and reduce mistakes of interpretation. They are also useful in providing cross references to related information.

Tenderers may request that their tender documentation be returned, for the sole purpose of reviewing the marginal notes made by evaluators (this is advised, for example by Silver [1990]). We see no reason why this practice should be permitted to degrade the evaluation activity. To this end, we recommend that tenderers be advised in the RFT that the tenders will eventually be destroyed, with certificates of destruction being provided to the tenderers.

6.5 Assessing risk

Performance risk is related to the impact on the mission objectives. If one or more primary mission objectives are likely to be seriously affected, the impact is assessed as high. If secondary mission objectives are seriously affected, or the effectiveness or efficiency of primary missions is significantly affected, the impact is assessed as medium. Low impact is assigned to situations where there is a minor performance degradation.

Individual assessors will often have difficulties in assessing risk, because they are not in a position to decide the impact on mission objectives. In many cases, they will need to be assisted in this area by moderators and TEWG leaders.

CEPMAN (Part 3, Chapter 16) provides a good overview of risk assessment. When evaluating the performance and technical aspects of a tender, however, further guidance is needed to the assessment of performance risk. CEPMAN identifies 4 types of risk: cost, schedule, management and performance. Typically project risks fall into a combination of these and other categories. For example, correction of problems resulting from performance risk will often impact cost and schedule.

Performance risk is caused by several factors which may be regarded as risks in their own right.

a. **Technical risk** occurs when there is some risk that the technical objectives may not be achievable, i.e. "pushing the state of the art". Technical risk can be reduced by the review of trials results, if available, and/or detailed analyses by objective specialists, such as DSTO. Although there may be a large amount of development required for the architecture and software of computer based systems, the performance sought and offered will normally be achievable, and presents a low technical risk.

b. **Development risk** is the risk that a developer cannot achieve the technical performance, even where there is no technical risk. In general this risk depends on the developer's experience and technical management, coupled with their commitment to meet contracted requirements. The only significant control that an acquirer can exert on development risk is in the selection of the supplier. The development risk associated with the development or major redevelopment of a tactical data system for a combat system, for example, will always be at least medium, usually resulting in an impact to schedule.

c. **Specification risk** (sometimes referred to as requirements risk) is widely accepted as the major cause of problems in the development of complex computer based systems. It is the risk that, although the delivered system may meet the letter of the requirements, it is not what the customer wanted or expected. It may be caused by errors, ambiguities or lack of detail in the specification, or simply by an inability or unwillingness (for commercial reasons) on the contractor's part to correctly interpret the specification. This latter situation will be aggravated by poorly specified requirements. Specification risk may be reduced by paying more attention to the development of specifications both prior to the RFT and during contract negotiations, and by the use of V&V activities during development.

d. **Contractual risk** is typically caused by inadequacies in the contract, providing insufficient authority or monitoring of development, as well as the acquirer being poorly supported in technical issues. It should be noted that a "good" contract will not automatically result in success. If the acquirer cannot take advantage of its rights under the contract, there is little value in its having those rights.

Performance risk (like most risks) is in most cases predictable and avoidable and in other cases controllable. The difficulty is in providing staff who have the appropriate talent, skills and experience at the correct stages of the project.

6.6 Moderation and collation of criteria

Moderation is an activity which helps ensure that the evaluation is objective and consistent with regard to its treatment of different tenders, and is consistent in

the different technical evaluation areas [CCTA 1990]. Moderators are experienced evaluators, often TEWG leaders, who should do the following:

- Ensure that the assessors of individual criteria in their area of responsibility are aware of what they should do and how they should do it.
- Review assessments in conjunction with the assessors, ensuring that the assessments are justified and fair.
- Ensure that the assessors have the necessary knowledge and experience to adequately assess the criteria, and if necessary take remedial action. (The assessors will, of course, have been initially selected for their ability to do their task, but different responses may have addressed technological areas outside an assessor's expertise.)
- Ensure that the same standards are being applied in all areas within their control.
- Consider and respond to requests from assessors for additional information. In some cases this will necessitate the clarification of information from the tenderers.

Collation involves assessing the node criteria, using as a basis the criteria below them in the tree. This process continues until assessments of the high level technical criteria are complete. Note that it is not essential that assessments be provided for all nodes. It will therefore be useful to identify the nodes which will be used for collation, early in the evaluation process. In Figure 2 for example, the Integration criterion might be omitted. Instead, the Fusion and Automation criteria could be collated into the Integrated solution criterion, along with the other level 6 criteria in that branch.

Where lower level criteria are compliant, the collation will follow the weights identified in the evaluation model as a guide to providing ratings. Where there are non-compliances, these may have a larger affect on the ratings at higher levels, depending on their impact. In some cases, for example, a low level non-compliance may not be regarded as serious enough to affect compliance at a higher level. The effect of such areas of non-compliance is not lost, however. All areas of non-compliance should be collected separately, and included in the final evaluation.

As part of the collation and moderation process, a narrative qualitative summary of significant capabilities and subsystems should be developed. Candidate areas from Figure 1 might include Combat data system, Tactical navigation and Radar, for example.

These interim reports should include a summary of each capability or subsystem offered (noting that the same equipment may be offered by more than one tenderer) including:

- A broad description of how the requirement is met.
- How widely used the proposed solution is.
- An assessment of the redevelopment of existing equipment required, if applicable.
- General comments on strengths, weaknesses, concerns and potential problems.

- Some indication of the comparative value of the equipment related to its rivals.

It should also include a preliminary ranking of proposals in this area, with justifications.

6.7 Writing the evaluation report

In large projects the technical evaluation will normally be presented in a separate evaluation report, parts of which will be used in the Source Evaluation Report (SER). In smaller projects the technical evaluation report may be included in the SER in its entirety. In either case, preparation of the technical evaluation report must be focused on the needs and guidance provided in CEPMAN or other high level guidance on preparing the SER (see section 1.4.1). Rather than addressing the total contents of the evaluation report, which can be found in CEPMAN, this section discusses how the technical rankings can be presented.

The technical evaluation report must clearly show the logic in reaching the final rankings and recommendations. This will include:

- Where several parts of the evaluation are drawn together to reach a conclusion, the weighting given to each aspect, or the relative priority of the aspects, must be shown.
- Where tenders are ranked against the capabilities, subsystems and the system itself, their relative strengths must be shown. This might be done using words such as "A is strongly preferred to B" or "B offers only marginal benefits over C", or even by using some form of numerical scoring system.

It is suggested that each major capability and subsystem evaluated should be addressed (using the evaluation database and the summaries discussed in section 6.6) with regard to compliance, performance and ranking. Other "system" features such as architecture, integration, interfaces, engineering and software should also be addressed. Significant non-compliances and their potential impact should also be discussed.

This information then needs to be considered against the major operational tasks to rank the systems for compliance and performance. Risks and the confidence level of the evaluation also need to be addressed (see section 6.5).

6.8 Resources

Apart from the more obvious "people" resources for assessment, moderation and the higher level evaluation activities, evaluations can result in quite different work patterns which should be recognised and catered for. One area where the resources must be available is in managing the evaluation tools. It is almost inevitable that changes will need to be made to the tools and to the environment in which the tool is used. If the tool is new, or severely modified from a previous evaluation, its first real test will be after the evaluation has started. Testing the tools with the expected number of users before this will be

difficult and is unlikely to be done effectively. There *will* be problems, and they will need to be fixed.

It is advisable that one or more personnel external to the TEWGs be designated for this role. Typical activities will include:

- Correcting defects or developing work arounds for the tools.
- Modifying the tools to adapt to changes in the evaluation model.
- Incorporating tool features that had not been anticipated.
- Backing up the evaluation database.
- Recovering from power and network failures.
- Forestalling or reacting to changes in the network configuration.
- Liaising with the local system administrators to ensure that the evaluation processing requirements are met, and are continuously available.

6.9 Providing information on the evaluation process to tenderers

There are several advantages in providing information to tenderers on how acquirers conduct their tender evaluations, i.e. in exposing the evaluation process to them. It should be noted that we are referring here to the process in general, not the evaluation model as developed for a specific project. The advantages are as follows:

- A well developed evaluation process will help tenderers to understand that the acquirer is being systematic in attempting to be fair and objective, and reduce complaints about the process.
- It will show them why information is requested, and how it will be used, allowing them to be more focused in the preparation of tenders, and in some cases reduce the amount of documentation tendered.
- It will show them the risks they are taking, and the penalties they may pay, in providing technical proposals which are incomplete, inconsistent incoherent, or misleading.

With regard to information specific to an individual project, CEPMAN requires project managers to provide a summary of the criteria which will be used in the evaluation (Part 4, Chapters 4 and 5). It is evident that this applies to the high level criteria, of which the extent of compliance with the technical and operational requirements is regarded as one criterion.

It will also be useful in some projects to provide some indication of the relative importance of high level technical criteria, possibly as a list in order of priority, but not providing detailed criteria or weighting factors [CCTA 1993; Rushforth et al. 1990]. Such information can provide the tenderers with an understanding of which system functions are more critical, and allow a more focused proposal.

6.10 Recording lessons learnt

One of the more critical activities in improving the evaluation process will be in recording the lessons learnt during each evaluation. It is important that evaluation personnel are aware that this information is being collected and a

point of contact for notifying specific problems and other relevant issues needs to be identified, so that comments can be made *during* the evaluation itself.

At the end of the evaluation, a proactive review of the activities as a whole needs to be made, possibly in conjunction with the use of questionnaires. Issues which need to be addressed should include the effectiveness and efficiency of:

- The specification as a basis for evaluation.
- The evaluation model.
- The RFT and tendered information.
- The overall evaluation process.

7. Commercially available tools

7.1 Evaluation database tools

We are aware of no commercially available tools which support an evaluation database to the extent required in complex projects. Tools marketed as evaluation tools are generally more involved with the scoring and final decision making processes, rather than the collection of the actual assessment details, or the collation of multiple ratings and textual information.

We consider that the tools we have seen are more applicable either for the evaluation of far simpler applications than the complex systems that are the subject of this study, or for decision making when there are major differences between how the competing offers address the problem. They are therefore seen as decision making tools, and are addressed in the next section.

7.2 Decision making tools

There are several tools which can be used in the decision making processes needed in evaluations. These may be used in the development of the evaluation model, in assigning weighting factors (see section 5.6) or in the final ranking of offers.

a. APET. APET (All Purpose Evaluation Tool) is a tool developed by the Consultancy Services Unit (CSU) of the Department of Finance. The CSU can provide consultative assistance in evaluations, and APET is supplied as part of the consultancy. APET uses the Analytical Hierarchical Process (AHP - see section 5.6) to combine numerically scored assessments into a final, numerically based decision, and is particularly strong in its graphical presentation of the results.

b. Select Gain. Select Gain is a method and tool based on the MAUT (Multiattribute Utility Theory) method. It is marketed by Planning Support Inc. who offer consultative services using this tool. Select Gain was developed by Dr. Edward Lewis, of the Computing Science Department at the Australian Defence Force Academy. Dr. Lewis also provides consulting services with regard to evaluations.

c. Criterium Decision Plus. Criterium is a relatively low priced tool (under $1000), produced by Sygenex and available from Technology Australasia. It includes functions supporting the AHP and MAUT methods and other decision making techniques. It also has "brainstorming" features which could be useful in the early development of the evaluation model. We reviewed version 1.1.

d. Data. Data is another low priced tool produced by TreeÄge and is also available from Technology Australasia. While not strictly aimed at evaluations it is designed to support decision making using decision trees, with an easy to use and highly visual user interface. We consider that it could be useful in the final stages of an evaluation, particularly in the incorporation of risk assessments. We reviewed version 2.6.

8. Requirements for an evaluation tool

8.1 Overview

This section discusses the broad requirements for a general purpose evaluation database tool for technical evaluations of complex systems and makes suggestions on how such a tool might be developed. It should be stressed that the requirements are not *defined* here, nor is a design solution. The needs for such a tool can only be assessed in detail by an analysis of the needs of all stakeholders, in conjunction with the committed involvement of typical users of the tool. Instead, the comments here are intended to be used as a basis for further investigation.

The problem of an assessor completing assessments is not considered a serious issue - this can be provided relatively easily using a low priced database tool. The main problems, and the areas in which we believe the most benefit can be gained is in the functions which address multiple forms in the evaluation, and the administration of the evaluation model.

It is therefore recommended that acquirers should consider the development of a tool which may form the basis for all technical evaluations. The remainder of this section contains suggestions regarding some of the functions that such a tool might encompass.

8.2 Terminology

The following establishes a terminology for some of the items and functions used in a typical tool.

Assessment form	A screen area designed for input or display of information relating to a single assessment.
Assessment record	Normally refers to the information relating to a single assessment, regardless of which database tables the different information may be contained in.

Configuration	A set of assessment records corresponding to all the criteria in the evaluation model which apply to a single proposed configuration.
Field	An individual data item in the database, and the equivalent area for entering or viewing the data item on a form or list. A separate field will normally be used for each rating, requirement number, security classification etc.
List	A displayed list of assessment information with each row corresponding to a single assessment record. A list will normally differ from an assessment form in that several records are shown and can be compared, but the number of different fields shown will be typically less than on a form.
Slice	A subset of the assessment records, providing a specific view of the evaluation.
Subsystem group	A set of assessment records which constitute an assessment of a single subsystem, which may be common to more than one configuration. It will usually, but not always, consist of a branch of the evaluation tree.

8.3 Structural issues - handling variants, options and common solutions

Variants, options and common solutions were discussed in section 5.8. The use of automated tools can make the handling of these issues much more difficult than in the older paper based methods.

The problems these issues raise is best illustrated by a typical example. Table 4 shows three subsystems of a combat system offered by three different tenderers. Subsystems shown in brackets are proposed options, and "var" indicates slight variations on a subsystem offered by another tenderer.

Table 4. Example of the effects of variants, options and common solutions

	Tender 1	Tender 2	Tender 3
Tactical data subsystem	TDS-A	TDS-B [TDS-C]	TDS-A var [TDS-B]
Radar	Radar-A [Radar-B]	Radar-B [Radar-A var] [Radar-C]	Radar-A
Gun	Gun-A	Gun-A [Gun-B]	Gun-A [Gun-B var]

There are actually only 8 different subsystems offered here, with additional variations for 3 of them. Assessors, having determined that the technical proposals are identical in these areas, should only need to undertake assessments for the 8, with minor additional assessments for the "var" offers.

The database representation of the options and common solutions shown in Table 4 is not straightforward for the following reasons.

- If all possible configurations (variants combining all options) need to be separately evaluated then there are 18 different configurations based on these subsystems alone: 2 for Tender 1, 12 for Tender 2 and 4 for Tender 3. Any other options in the tenders will further increase these numbers. Clearly all options cannot be handled by creating different configurations.

- Options cannot always be isolated to a single branch (as discussed in section 5.8) or assessed in isolation. The tactical data system (TDS) is the heart of a combat systems with interfaces to most other subsystems. Optional TDSs are one good case for considering the application of different configurations.
- Assessments of the same equipment or software offered in more than one tender should not be duplicated as separate assessment records. Assessments are often changed during the evaluation (during moderation, for example) and duplication will result in the risk of some duplicates not being updated.
- The variant options shown in the example may only result in the changing of the assessments of one or two criteria. Duplication should also be avoided in this situation.

The solution to this problem must allow for configurations to use common assessments, i.e. each assessment record may apply to a number of different configurations, including those in different tenders. The solution must also be flexible, because the actual criteria affected will not be completely known until the detailed assessments begin. Decisions on what assessments apply to which configurations will need to be made by those overseeing the technical evaluation, not by individual assessors, partly because of the need to ensure that the solutions offered are indeed identical. This is a complex and ongoing task.

To prevent assessment errors, the offers and options to which each assessment applies must be evident from looking at the assessment form. Similarly it must be possible to view, and provide reports on, all possible configurations (with options as selected by the tool user).

8.4 Functional requirements

These requirements are suggested as some of the needs for an evaluation database tool.

a. Compatibility with specification development tools. It is important that the evaluation tool is compatible with the specification development tool which may be selected by the acquirer. Much of the information collected in a requirements database will be relevant during the evaluation.

b. User interface. The user interface needs to be defined to minimise the effort and training needed to operate the tool. It should be based on the needs and experience of the users rather than what is easy for the developer. Semi-graphical methods should be used for navigating through and selecting parts of the evaluation tree, and displaying and modifying configurations and options, for example.

c. Selecting slices. Functions for users to select a slice of the total assessment records using a combination of search fields, including:

    - Tenderer(s), configuration(s) and/or subsystem(s).
    - Leaf or node assessments.

- Security classification.
- Branch (all nodes and leaves below a particular node).
- Different ratings or whether text has been entered in particular fields.

Functions to expand or contract the currently selected slice and to save and later reinitiate a series of different slices.

Functions to sort the list on any of, or a combination of, the above fields.

d. Lists. Functions to display a slice in a list, with the format (fields, field widths and order of display) configurable by the user.

Functions to change from a list view to the equivalent form for the record indicated, and vice versa.

e. Field copying. Functions to copy a field or fields in a record to the other records, or to other selected records, in a slice. This will be necessary when, for example, a set of forms are likely to contain the same, or very similar information regarding the references of the evidence for the assessment.

f. Multiple views. The ability to view more than one slice at a time. This will be used when it is necessary to refer to other assessments, while maintaining the current view.

g. Integrity checking. Functions to check the integrity of assessment records and configurations. Typical checking will include:

- Checking that compulsory fields are completed, and the entered fields are consistent (e.g. that if risk is identified, the impact is entered). This type of checking should be controllable by the user - it should not be necessary to complete an assessment before leaving the record, but all completed forms should have been checked for consistency.
- Checking that each configuration has all the required assessment records included.

h. Administrative functions. Functions to facilitate the splitting, merging and reassignment of assessment records to cater for variants, options and common solutions.

i. Progress reports. Functions to provide predetermined and user defined progress reports, both on screen and printer, for the current slice. Progress reports will typically address the status of the records in the slice. Typical information will include:

- Number of assessments completed.
- Number of completed assessments above or below a given rating for compliance, confidence and risk.
- Distribution by tender, configuration, subsystem etc.

- A comparison between the current results for different tenders, showing, for example, the differences in level of compliance and perceived risk between tenders.

j. Evaluation reports. Functions to print selected information of a particular slice of the database.

8.5 Development approach

Whether an evaluation tool is developed in-house or commercially, it will still be necessary to determine the detailed requirements for such a tool. In our experience, the main deficiencies in tools stem from the following causes:

- Underestimation of the needs of the evaluation, and overestimation of the power of the tools used.
- Insufficient consideration being paid to the exceptional circumstances which occur in assessments.
- Lack of usability testing across a range of experienced staff with evaluation experience.
- Inadequate consideration of the complexities caused by variants, options and common solutions.

It is therefore recommended that, if a decision is made to develop such a tool, these issues are addressed by a team of staff with diverse and extensive experience in evaluations, prior to the selection or design of a tool.

9. Conclusions and recommendations

The technical evaluation of complex systems forms the basis of some of the most important decisions made in the procurement of complex military systems. It is a time consuming and intensive activity which can be highly stressful for the personnel participating.

In our opinion, both the effectiveness and efficiency of technical evaluations can be significantly improved by the application of appropriate processes and tools, and by improvement of the process based on the lessons learnt from evaluation activities. This report provides suggestions to help improve the evaluation process and to aid the selection of appropriate tools.

Some specific recommendations are as follows:

- Establish a defined and monitored evaluation process based on the findings and recommendations of this study. The process should then be monitored and improvements made on the basis of lessons learnt in evaluations.
- Consider the development of a general purpose evaluation database tool to be used in technical and other evaluation areas.
- Plan and provide dedicated resources for the management of evaluation tools during evaluations. A large amount of effort is required for this task, and if not provided, the efficiency of the evaluation process will be degraded.

- Emphasis be placed on providing guidance to assessors on how to assess tenders against evaluation criteria. Assessors need comprehensive and authoritative guidance in order to effectively and efficiently carry out their tasks.
- Guidelines be established with regard to advice to tenderers on the content and format of their technical proposals. The benefits are likely to be reduced quantity of documentation and increased quality of the tenders.
- Use numerically based evaluation methods only as a check against qualitative methods. We consider that they are unsuitable for the evaluation of complex systems but can sometimes be useful as a secondary method to check the results obtained from qualitative methods.

10. References

1994, *Costs of tendering industry survey*, Department of Defence, Australian Government Publishing Service, Canberra.

CCTA 1990, *Overview and procedures volume*, CCTA ISE appraisal and evaluation library, HMSO, London.

CCTA 1993, *Producing a Statement of Service Requirements*, CCTA Market testing IS/IT series, HMSO, London.

CEPMAN 1995, *The capital equipment procurement manual* AL8.

CPG 1 1989, *Getting value for money*, Commonwealth Procurement Guideline No. 1, Department of Administrative Services, Purchasing Reform Group, Australian Government Publishing Service, Canberra.

CPG 8 1992, *Managing risk in procurement*, Commonwealth Procurement Guideline No. 1, Department of Administrative Services, Office for Better Buying, Australian Government Publishing Service, Canberra.

Gabb, Andrew P. and Derek E. Henderson 1995a, *Navy specification study: Report 1 - Industry survey*, Technical Report DSTO-TR-0190, DSTO.

Gabb, Andrew P. and Derek E. Henderson 1995b, *Navy specification study: Report 2 - Current policy and practice*, Technical Report DSTO-TR-0191, DSTO.

Gabb, Andrew P. and Derek E. Henderson 1995c, *Navy specification study: Report 3 - Requirements and specifications*, Technical Report DSTO-TR-0192, DSTO.

Gabb, Andrew P. and Derek E. Henderson 1995d, *Navy specification study: Report 4 - Executive summary and final recommendations*, Technical Report DSTO-TR-0193, DSTO.

Gabb, Andrew P. and Derek E. Henderson 1995e, *A review of Navy's technical and operational evaluation practices*, Technical Report DSTO-TR-0194, DSTO.

Macphee, Ian 1992, *Independent review of the Civil Aviation Authority's tender evaluation process for the Australian Advanced Air Traffic System*, Australian Government Publishing Service, Canberra.

Millett, Jack 1994, *Guide for the preparation and use of performance specifications*, AMC-P 715-17, U.S. Army Material Command, Alexandria, Virginia, March .

Roetzheim, William H. 1991, *Developing software to government standards*, Prentice Hall, Englewood Cliffs, New Jersey.

Rushforth, Shelagh, Richard L. Davis and Linsay Gallon 1990, *Review of procurement on behalf of the Procurement Division*, Central Computer and Telecommunications Agency (UK) Study performed by Electronics Facilities Design, EFD/8931/FR01A, January.

Saaty, Thomas L. 1980, *The analytical hierarchical process*, McGraw-Hill, New York.

Silver, Hyman 1990, *Creating the superior technical proposal*, training videotape, Silver Crown Productions, Los Angeles.

Tullous, Raydel and Utecht, Richard L. 1994, "A decision support system for integration of vendor selection task", *Journal of applied business research*, Volume 10, Number 1.

DISTRIBUTION

Technical and Operational Tender Evaluations for Complex Military Systems

Andrew P. Gabb and Derek E. Henderson

	Number of Copies
DEPARTMENT OF DEFENCE	
Defence Science and Technology Organisation	
Chief Defence Scientist and members of the	) 1 shared copy
DSTO Central Office Executive	) for circulation
Counsellor, Defence Science, London	(Document Control sheet)
Counsellor, Defence Science, Washington	(Document Control sheet)
Senior Defence Scientific Adviser	) 1 shared copy
Scientific Adviser - POLCOM	)
Assistant Secretary Science Industry Interaction	1
Director, Aeronautical & Maritime Research Laboratory	1
Navy Scientific Adviser (NSA)	1
Scientific Adviser, Army (SA-A)	1
Air Force Scientific Adviser (AFSA)	1
Electronics and Surveillance Research Laboratory	
Chief Information Technology Division	1
Research Leader Command & Control and Intelligence Systems	1
Research Leader Command, Control and Communications	1
Research Leader Military Computing Systems	1
Executive Officer, Information Technology Division	(Document Control sheet)
Manager, Human Systems Integration Group	(Document Control sheet)
Head, Software Engineering Group	(Document Control sheet)
Head, Trusted Computer Systems Group	(Document Control sheet)
Head, Advanced Computer Capabilities Group	(Document Control sheet)
Head, Intelligence Systems Group	(Document Control sheet)
Head, Systems Simulation and Assessment Group	(Document Control sheet)
Head, Exercise Analysis Group	(Document Control sheet)
Head, Computer Systems Architecture Group	(Document Control sheet)
Head Command Support Systems Group	(Document Control sheet)
Head Information Management and Fusion Group	(Document Control sheet)
Head, C3I Systems Engineering Group	1
Mr Andrew Gabb, C3ISE Group, ITD	15
Mr Derek Henderson, C3ISE Group, ITD	3
Publications and Publicity Officer, ITD	1
Navy Office	
Director, Naval Combat Systems Engineering	5

Libraries and Information Services

Defence Central Library - Technical Reports Centre	1
Manager Document Exchange Centre (MDEC) (for retention)	1
Additional copies which are to be sent through MDEC DIS for distribution:	
National Technical Information Centre, United States	2
Defence Research Information Centre, United Kingdom	2
Director Scientific Information Services, Canada	1
Ministry of Defence, New Zealand	1
National Library of Australia	1
Defence Science and Technology Organisation Salisbury, Research Library	2
Library Defence Signals Directorate Canberra	1
AGPS	1
British Library Document Supply Centre	1
Parliamentary Library of South Australia	1
The State Library of South Australia	1

Spares

Defence Science and Technology Organisation Salisbury, Research Library	6

Department of Defence

DOCUMENT CONTROL DATA SHEET

1. Page Classification	
UNCLASSIFIED	
2. Privacy Marking/Caveat	
N/A	

3a. AR Number	3b. Establishment No.	3c. Type of Report	4. Task Number	
AR-009-625	DSTO-TR-0303	TECHNICAL REPORT	NAV93/067	
5. Document Date	6. Cost Code	7. Security Classification	8. No. of Pages	54
April 1996	WA807716	[U] [U] [U] Document Title Abstract	9. No. of Refs.	18

10. Title

TECHNICAL AND OPERATIONAL TENDER EVALUATIONS FOR COMPLEX MILITARY SYSTEMS

S (Secret C (Confi) R (Rest) U (Unclass)

* For UNCLASSIFIED docs with a secondary distribution LIMITATION, use (L) in document box.

11. Author(s)

Andrew P. Gabb

Derek E. Henderson

12. Downgrading/ Delimiting Instructions

N/A

13a. Corporate Author and Address

Information Technology Division
Electronics and Surveillance Research Laboratory
PO Box 1500 SALISBURY SA 5108

14. Officer/Position responsible for

Security	SOESRL
Downgrading	CITD
Approval for release	CITD

13b. Task Sponsor

NAVY

15. Secondary Release Statement of this Document

APPROVED FOR PUBLIC RELEASE

Any enquiries outside stated limitations should be referred through DSTIC, Defence Information Services, Department of Defence, Anzac Park West, Canberra, ACT 2600.

16a. Deliberate Announcement

No limitation

16b. Casual Announcement (for citation in other documents)

[X] No Limitation | [ ] Ref. by Author, Doc No and date only

17. DEFTEST Descriptors

Computer Systems, Equipment Specifications, Tenders, Evaluation

18. DISCAT Subject Codes

19. Abstract

This report examines technical and operational tender evaluations for complex computer based military systems. Detailed recommendations are made with respect to getting the right technical proposals, developing an evaluation model and controlling the evaluation process.